

ROTATION PERIODS OF 34,030 KEPLER MAIN-SEQUENCE STARS: THE FULL AUTOCORRELATION SAMPLE

A. MCQUILLAN, T. MAZEH

School of Physics and Astronomy, Raymond and Beverly Sackler, Faculty of Exact Sciences, Tel Aviv University, 69978, Tel Aviv, Israel

S. AIGRAIN

Department of Physics, University of Oxford, Oxford, OX1 3RH, UK

Accepted ApJS 20th Feb 2014

ABSTRACT

We analyzed 3 years of data from the *Kepler* space mission to derive rotation periods of main-sequence stars below 6500 K. Our automated autocorrelation-based method detected rotation periods between 0.2 and 70 days for 34,030 (25.6%) of the 133,030 main-sequence *Kepler* targets (excluding known eclipsing binaries and Kepler Objects of Interest), making this the largest sample of stellar rotation periods to date. In this paper we consider the detailed features of the now well-populated period-temperature distribution and demonstrate that the period bimodality, first seen by McQuillan, Aigrain & Mazeh (2013) in the M-dwarf sample, persists to higher masses, becoming less visible above $0.6 M_{\odot}$. We show that these results are globally consistent with the existing ground-based rotation-period data and find that the upper envelope of the period distribution is broadly consistent with a gyrochronological age of 4.5 Gyrs, based on the isochrones of Barnes (2007), Mamajek & Hillenbrand (2008) and Meibom et al. (2009). We also performed a detailed comparison of our results to those of Reinhold et al. (2013) and Nielsen et al. (2013), who have measured rotation periods of field stars observed by *Kepler*. We examined the amplitude of periodic variability for the stars with detected rotation periods, and found a typical range between ~ 950 ppm (5th percentile) and $\sim 22,700$ ppm (95th percentile), with a median of $\sim 5,600$ ppm. We found typically higher amplitudes for shorter periods and lower effective temperatures, with an excess of low-amplitude stars above ~ 5400 K.

Subject headings: stars: rotation – methods: observational

1. INTRODUCTION

Throughout their main-sequence lifetime, stars lose angular momentum and spin down. For intermediate- and low-mass stars, this angular momentum loss is thought to occur via a magnetized wind that is linked to the stellar outer convection zone (e.g., Kawaler 1988; Bouvier et al. 1997). Therefore, the present-day rotation period reflects the integrated angular momentum loss history of the star.

It has been claimed that for main-sequence stars, the present day rotation period does not depend on the initial conditions, and therefore a tight relationship exists between the period, age and mass, which is known as gyrochronology, for which a range of empirical relations have been derived (e.g., Kawaler 1989; Barnes 2003, 2007, 2010; Mamajek & Hillenbrand 2008). In recent years there has also been considerable work towards understanding and modeling the underlying physics behind gyrochronology (Reiners & Mohanty 2012; Matt et al. 2012; Gallet & Bouvier 2013, and references therein). Therefore, deriving the rotation period for a large number of main-sequence stars has been a long-standing goal in stellar astrophysics, with the potential to shed light on stellar, and even planetary system angular-momentum evolution on one hand, and provide a new method to probe Galactic star formation history and structure on the other hand.

In recent years, the study of stellar rotation has seen

significant advances thanks to new observational programs and techniques. Previously, measurements of stellar rotation were typically performed using spectroscopy, via the rotational broadening of absorption lines (e.g., Kaler 1989). However, this technique yielded only model-dependent constraints on the rotation rate due to the dependence on stellar radius and inclination, and were limited to relatively fast rotators.

The era of exoplanet transit surveys provided a new approach to stellar rotation studies, making it possible to measure the rotation period directly for a large number of stars simultaneously. This is done by detecting quasi-periodic brightness variations in the time series photometry, caused by magnetically active regions repeatedly crossing the visible hemisphere as the star rotates (e.g., Irwin et al. 2009).

Open cluster surveys over the past decade have provided thousands of rotation periods for low-mass stars with ages up to ~ 650 Myr, an overview of which can be found in Irwin & Bouvier (2009) and Bouvier (2013). More recent additions to the literature on rotation of young low-mass stars include studies of M35 (310 periods) (Meibom et al. 2009), Coma Berenices (Collier Cameron et al. 2009), M34 (83 periods) (Meibom et al. 2011), and h Persei (586) Moraux et al. (2013). Together, these data have provided a relatively complete, but complex picture of pre-main-sequence rotational evolution.

Until recently, rotation periods for older stars remained scarce, since their slow rotation and low-amplitude mod-

ulation make them difficult to detect with either spectroscopy or ground-based photometry. Notable exceptions include 112 main sequence F, G and K stars observed as part of the Mount Wilson H-K project (Baliunas et al. 1996, and references therein), 1727 mid-F to mid-K stars observed by the *CoRoT* satellite (Affer et al. 2012), 41 low-mass ($0.1\text{--}0.3 M_{\odot}$) stars from the *MEarth* survey (Irwin et al. 2011), and ~ 2000 field K and M stars observed by the HATNet survey (Hartman et al. 2011). The periods of Affer et al. (2012) are based on up to 5 months of CoRoT data and yet include a significant fraction of periods greater than 60 days. Since many of these periods are more than half the length of the dataset, they should be treated with caution and we opted to omit these results from the comparison section of this paper.

Asteroseismology studies performed on a small subset of *Kepler* and CoRoT targets also provide an insight into stellar rotation, since the oscillation spectrum of a star can reveal information about surface and interior rotation rates (e.g., Gizon 2002; Gizon & Solanki 2004). For a recent review on observational studies of stellar rotation, see Bouvier (2013).

Data from the Kepler space mission (Borucki et al. 2010; Koch et al. 2010) are revolutionizing the study of stellar rotation, providing 4 yrs of almost uninterrupted photometry with an unprecedented level of precision and time sampling, for a very large sample of stars. This allows rotation periods to be detected for slowly rotating stars with low-amplitude modulation, for a well-defined large sample.

Rotation studies on subsets of the *Kepler* data include the measurement of 265 rotation periods for stars with $T_{\text{eff}} \leq 5200$ K and $\log g \geq 4.0$ dex observed by *Kepler* for 1–2 quarters through the Cycle 1 Guest Observer program (Harrison et al. 2012), 1570 M-dwarf rotation periods using quarters 1–4 (Q1–4) (McQuillan et al. 2013a), 737 of the Kepler Objects of Interest (KOIs) using Q3–14 data (McQuillan et al. 2013b), ~ 950 of the KOIs using Q9 (Walkowicz & Basri 2013).

Two previous studies focussing on the broader *Kepler* sample are those of Reinhold et al. (2013), with an emphasis on differential rotation, who derive $\sim 24,000$ periods using Q3, and Nielsen et al. (2013), who measured $\sim 12,000$ periods from Q2–Q9, and compare to previous spectroscopic studies. These studies, discussed in Section 7, use Fourier-based detection methods and conservative automatic selection criteria.

In this paper we describe the design and application of AutoACF, an automated version of the autocorrelation function (ACF) method introduced by McQuillan et al. (2013a) for stellar rotation period measurement. Using AutoACF we derived 34,030 rotation periods from 133,030 main-sequence targets observed by *Kepler*. We outline the sample selection in Section 2 and the period detection method in Section 3. In Section 4 we present our results, and examine the mass-period and temperature-period distributions in Section 4.1, and the amplitude of periodic variability in Section 4.2. In Section 5 we compare our results to empirical gyrochronology relations, and in Section 6 we compare our results to those of Reinhold et al. (2013) and Nielsen et al. (2013). We summarize our results in Section 7 and provide full details of the ACF automation method in the Appendix.

2. SAMPLE SELECTION

This analysis made use of the public releases 14–19 of quarter 3–14 (Q3–Q14) light curves (LCs), which were downloaded from the Kepler mission archive¹. The data of Q0 and Q1 were omitted due to their short duration, and of Q2 due to significant residual systematics. We used the data corrected for instrumental systematics using PDC-MAP (Smith et al. 2012; Stumpe et al. 2012), which removes the majority of instrumental glitches and systematic trends using a Bayesian approach, while retaining most real variability.

The initial list of all targets observed by *Kepler* contains $\sim 195,000$ stars. To select only main-sequence stars, we used the $T_{\text{eff}}\text{--}\log g$ and color–color cuts advocated by Ciardi et al. (2011), which remove $\sim 32,000$ giants from the *Kepler* sample. At this stage we also excluded ~ 2000 targets without T_{eff} or $\log g$ values. Where available for the low-mass stars, we use the improved values of Dressing & Charbonneau (2013) in place of the KIC parameters. A more recent version of the KIC is available (Huber et al. 2014), however, we have opted to use the original since this provides the most homogeneous parameters. We checked that using the updated KIC values does not alter the results of this paper.

Since eclipsing binary (EB) and planetary transit signals may affect the detection of rotation periods, we removed the 2611 known EBs and 4799 KOIs from the sample. The EB list was downloaded from the Villanova Eclipsing Binary web site² which has an up-to-date and extended version of the *Kepler* EB catalogs of Prša et al. (2011) and Slawson et al. (2011). The EB catalog used in this work was downloaded on 7th Sept 2013. The KOI list was downloaded from the NASA Exoplanet Archive³ (Akeson et al. 2013) on 8th Sept 2013 and includes confirmed planets, candidates, false positives and targets which have not been assigned a classification (labelled ‘not dispositioned’).

At this stage 154,925 targets remained in our list. Some targets were not observed in all quarters, and hence we selected only targets observed in at least 8 out of the 12 quarters used in this study, leaving a sample of 139,638 stars. Finally, since we are interested only in the rotation periods of stars with convective envelopes which spin down during their lifetimes, we introduce a limit of $T_{\text{eff}} < 6500$ K. This also greatly reduces the potential for contamination of the sample by pulsating stars from the instability strip. The final sample on which the period search was run consisted of 133,030 stars.

3. PERIOD DETECTION

Periodic detection was performed using the ACF method, described in detail in McQuillan, Aigrain & Mazeh (2013). The ACF measures the degree of self-similarity of an LC over a range of time lags. In the case of rotational modulation, repeated spot crossings over the visible stellar hemisphere produce ACF peaks at lags corresponding to the stellar rotation period and its integer multiples.

¹ <http://archive.stsci.edu/kepler>

² <http://keplerebs.villanova.edu>

³ <http://exoplanetarchive.ipac.caltech.edu>

The ACF has been shown to produce clear and robust results even when the amplitude and phase of the photometric modulation evolve significantly, and when systematic effects and long-term trends are present (McQuillan et al. 2013a). For this reason, we adopt the ACF method of period detection over the more widely used Fourier-based methods, which can be strongly affected by residual systematics. The clarity of output and the robust nature of the ACF make it amenable to automation, without requiring such conservative thresholds as the Fourier-based approaches. See McQuillan et al. (2013a) for a detailed comparison of the ACF method with Fourier-based approaches.

We preprocessed the LCs following the method of McQuillan et al. (2013a), in which we median normalized each ~ 90 day quarter of *Kepler* data. In an update to the first application of this code, we now map the data to an evenly spaced grid of 29.4 minutes to match the *Kepler* long cadence, and fill any missing flux values (for example in the gaps between quarters) with zeros to prevent them contributing to the ACF.

We then computed the ACF for each LC and identified the peak corresponding to the rotation period, which is followed by series of peaks at lags equal to integer multiples of the rotation period. We then used the locations of this set of peaks to determine the exact period of the target LC. In a minor modification to the original code, we defined the period as the gradient of a straight-line fit to the ACF peak positions, as a function of peak number. Up to 4 consecutive peaks at near integer multiples of the selected ACF peak were used, as well as the intercept at (0,0). This provides a more reliable period measurement and uncertainty than the median and scatter of the peaks, which were used by McQuillan et al. (2013a).

For stars where active regions are present on opposite hemispheres of the star, a partial correlation exists at half the rotation period, leading to additional peaks in the ACF. If the second peak in the ACF is larger than those on either side, the ACF code identifies the larger peak (and integer multiples of it) as presenting the correct rotation period. Of the 5567 cases for which the ACF selected the second peak, we visually checked all cases for which the peak height difference between the first and the second peak was less than 20%, and corrected any erroneous peak selections.

McQuillan et al. (2013a) focussed on a small sample of 2483 cool *Kepler* targets, and visually examined all LCs to verify each rotation period detection by the ACF. Since this is not feasible for the 133,030 stars in our sample, we have automated the selection procedure (AutoACF). Using a visually verified training set with 1000 confirmed periodic targets per 500 K T_{eff} bin, we sought the parameters of the ACF which separate likely true period detections from false positives originating from noise or instrumental systematics.

We performed the selection in two stages, first testing the consistency of the ACF detection in separate data sections, and then the significance of the detection. A good criterion for real astrophysical periodicity is that it can be detected in multiple regions of the LC, since ACF peaks caused by systematics or artifacts are less likely to appear in different section of the LC. Therefore, we required the period found for the whole Q3–Q14 LC to

be detected also in 2 or more individual segments, where a segment is defined as 3 consecutive *Kepler* quarters (Q3–Q5, Q6–Q8, Q9–Q11, Q12–Q14).

To identify real periods based on the detection significance, we used the height of the primary ACF peak with respect to the troughs on each side, which we call ‘local peak height’ (LPH). Clear period detections have higher LPHs than those suspected to originate from noise. We assigned a weighting, w , to each period detection according to its LPH and position in T_{eff} -LPH-Period space. A threshold w_{thres} , can then be selected to obtain the desired compromise between real detections and false positives. For this work we selected $w_{\text{thres}} = 0.25$. We tested that altering the input parameters and functions used in AutoACF, and the threshold weight, w_{thres} , does not significantly alter the resulting period distribution. Full details of the selection method are given in Appendix A.

The first application of AutoACF derived rotation periods for 33,694 stars. We then compared our results to those of Reinhold et al. (2013) and Nielsen et al. (2013) who have also performed rotation period studies of the *Kepler* sample. This comparison, described in detail in Section 6.3, identified a subset of stars for which both other works were able to detect rotation periods and AutoACF was not. These missed detections had two dominant origins: faults in the LC, and multiple periods. We implemented an additional level of processing to the LCs and ACFs of stars where no rotation period had been determined to detect periods initially missed due to faults (marked as ‘BQR’ in Table 1) and multiple periods (marked as ‘SM1’ and ‘SM2’ in Table 1). Details of this procedure are described in Appendix B. These additions bring the total number of rotation period detections to 34,030. Differences between the results of McQuillan et al. (2013a) and the current study are minimal and dominated by the improvement in detection from the longer baseline, rather than the difference between the methods.

4. RESULTS

Using AutoACF we derived rotation periods for 34,030 of the stars in the *Kepler* sample defined in Section 2. These results are shown in Table 1.

From the application of AutoACF to the visually examined training set, we estimate that $> 95\%$ of the derived periods represent true rotation periods. Of the remaining $\sim 5\%$, half are probable rotation period detections and half are false detections triggered by systematics or likely pulsators (based on visual examination of the LC and ACF). Based on the application of AutoACF to the training sample, we estimate that an additional ~ 5000 periods could have been detected by eye, and that these are predominantly long-period or low-amplitude signals. It may be possible to augment the periodic sample by visual examination of targets falling just below the selected w_{thres} value.

Table 2 shows the subset of stars to which we applied AutoACF but the resulting detection did not meet our criteria. Table 3 shows the periodic detection fraction per 500 K T_{eff} bin. This fraction goes from ~ 0.8 for the coolest stars, with temperatures below 4000 K, to ~ 0.2 around 6000 K.

TABLE 1

THE ROTATION PERIOD MEASUREMENTS FOR THE 34,030 STARS PRESENTED IN THIS WORK. THIS TABLE IS AVAILABLE IN ITS ENTIRETY IN A MACHINE-READABLE FORM IN THE ONLINE SUPPLEMENTARY MATERIAL AND AT www.astro.tau.ac.il/~amy. A PORTION IS SHOWN HERE FOR GUIDANCE REGARDING ITS FORM AND CONTENT. T_{eff} AND $\log g$ ARE FROM THE KIC OR, WHERE AVAILABLE, FROM DRESSING & CHARBONNEAU (2013), INDICATED BY A DC VALUE OF 1. THE MASS, M , WAS DERIVED FROM T_{eff} USING THE ISOCHRONES OF BARAFFE ET AL. (1998). THE REMAINING COLUMNS ARE: PERIOD, P_{rot} , AND PERIOD ERROR, σ_P ; THE AVERAGE AMPLITUDE OF VARIABILITY WITHIN ONE PERIOD, R_{per} ; THE LOCAL PEAK HEIGHT, LPH; AND THE ASSIGNED WEIGHT, w . THE FLAG COLUMN INDICATES WHETHER FAULT CORRECTION ('BQR'), SOFT SMOOTHING ('SM1') OR HARD SMOOTHING ('SM2') WERE REQUIRED TO DETECT A PERIOD (SEE APPENDIX A AND B FOR FURTHER DETAILS).

KIC	T_{eff} (K)	$\log g$ (dex)	M (M_{\odot})	P_{rot} (days)	σ_P (days)	R_{per} (ppm)	LPH	w	DC	Flag
892376	3813	4.47	0.4699	1.532	0.007	7306.69	0.823	0.4503	0	—
1026146	4261	4.57	0.6472	14.891	0.120	11742.56	1.405	0.7229	0	—
1026474	4122	4.56	0.5914	1.569	0.006	30471.80	1.204	0.6061	0	—
1162635	3760	4.77	0.4497	15.678	0.019	10207.47	0.978	0.5445	1	—
1164102	4045	4.62	0.5606	31.496	0.474	5139.74	0.568	0.3939	0	—

TABLE 2

DETAILS OF THE 99,000 STARS WITH NO SIGNIFICANT PERIOD DETECTION. THIS TABLE IS AVAILABLE IN ITS ENTIRETY, IN A MACHINE-READABLE FORM IN THE ONLINE SUPPLEMENTARY MATERIAL AND AT www.astro.tau.ac.il/~amy. A PORTION IS SHOWN HERE FOR GUIDANCE REGARDING ITS FORM AND CONTENT. COLUMN DESCRIPTIONS ARE THE SAME AS FOR TABLE 1. TARGETS WITHOUT A w VALUE WERE REJECTED AT SELECTION PROCESS STAGE 1 BECAUSE THE PERIOD DETECTION DID NOT OCCUR IN ENOUGH SEGMENTS (SEE APPENDIX A FOR DETAILS). IN THESE CASES, P_{rot} , σ_P , LPH AND w ARE MARKED AS 'NAN'.

KIC	T_{eff} (K)	$\log g$ (dex)	M (M_{\odot})	P_{rot} (days)	σ_P (days)	LPH	w	DC
893305	4133	4.58	0.5958	nan	nan	nan	nan	0
1027110	4155	4.50	0.6046	1.701	0.039	0.299	0.1439	0
1027277	4326	4.57	0.6735	60.136	0.691	0.315	0.0876	0
1160660	4232	4.59	0.6355	nan	nan	nan	nan	0
1160684	3952	4.48	0.5239	0.419	0.090	0.150	0.0266	0

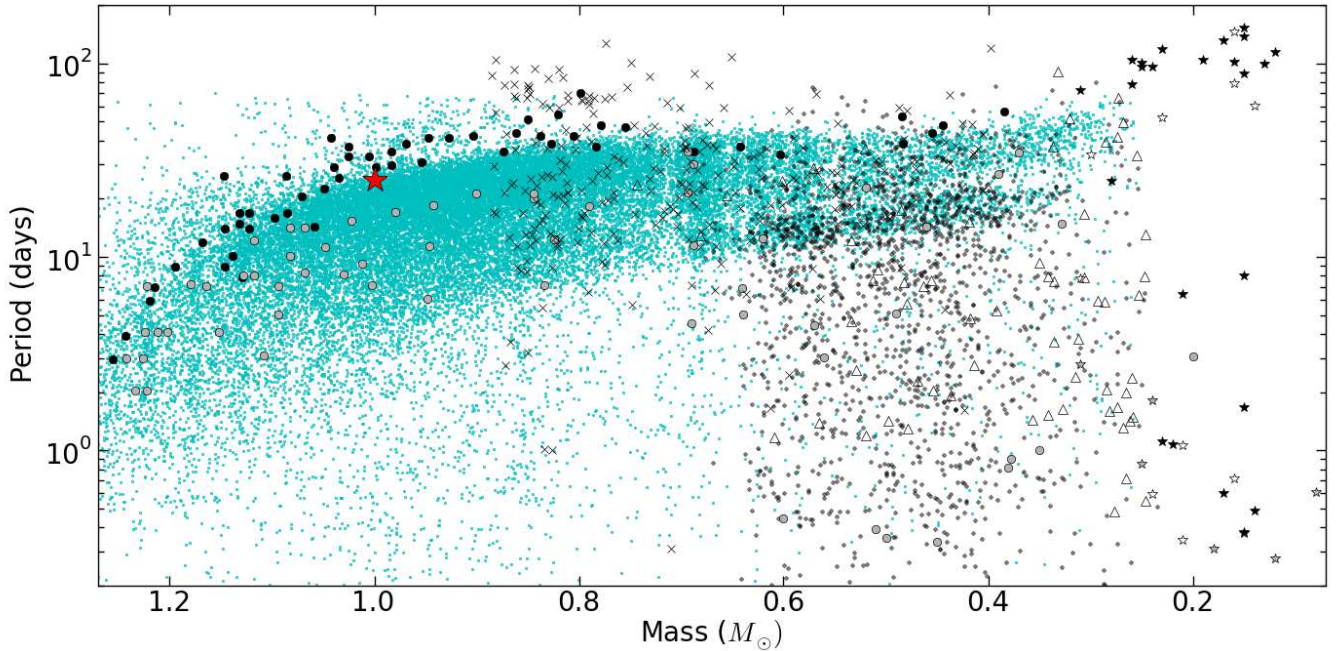


FIG. 1.— Period versus mass with comparison to previous rotation period measurements. The 34,030 new rotation periods derived using AutoACF are shown as cyan points. The mass was derived using the models of Baraffe et al. (1998), as described in the text. This figure also displays periods from Baliunas et al. (1996) and Kiraga & Stepien (2007) (114 circles) and *MEarth* data from Irwin et al. (2011) (41 stars), with grey and black symbols representing objects with young and old disk kinematics, respectively, all of which have available mass estimates. Additional M-dwarf periods from the WFCAM Transit Survey (Goulding et al. 2012), for which no kinematic classification is available (65 triangles), with masses derived from Mamajek (2011). Also included are periods from Hartman et al. (2011) (1686 small black dots), with mass estimates obtained using T_{eff} and the models of Baraffe et al. (1998), and periods from Harrison et al. (2012) (265 crosses), with masses derived from a $J - K$ to T_{eff} conversion using data from Kenyon & Hartmann (1995), and the isochrones of Baraffe et al. (1998).

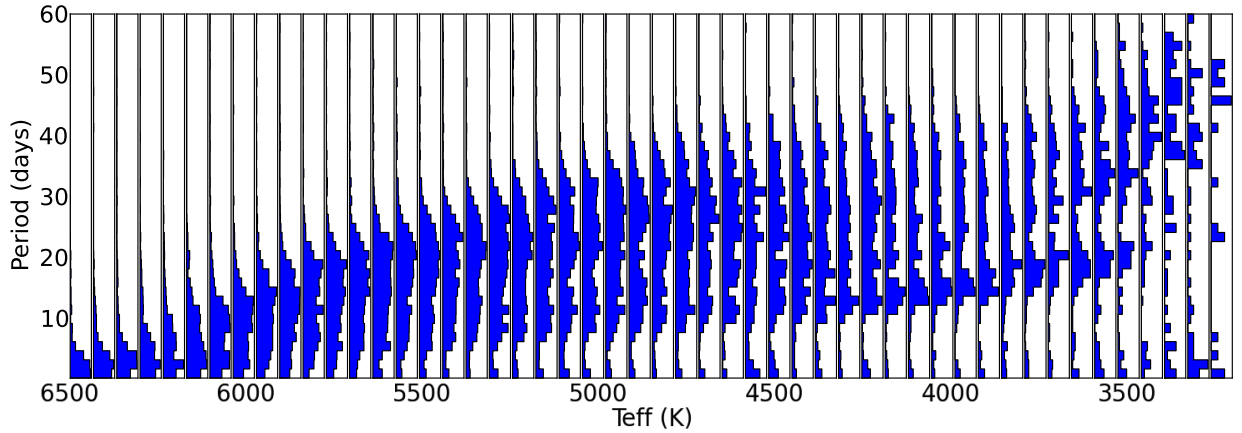


FIG. 2.— Period versus T_{eff} shown in histogram form for each temperature bin, clearly showing the bimodality in rotation period for cooler stars, and the slight dip in the upper envelope of the period distribution at ~ 4000 K.

TABLE 3
PERIODIC FRACTIONS ACROSS THE TEMPERATURE RANGE
EXAMINED.

T_{eff} (K)	Periodic Fraction
< 4000	0.83
4000–4500	0.69
4500–5000	0.43
5000–5500	0.27
5500–6000	0.16
6000–6500	0.20

4.1. Mass-Period and Temperature-Period Distributions

Figure 1 shows the mass-period distribution of the 34,030 stars with measured rotation periods, together with period derivations from previous work, most of which originate from ground-based observations. Mass, M , is calculated from the KIC T_{eff} using the stellar evolution models of Baraffe et al. (1998), using isochrone no. 1 for $M < 0.7 M_{\odot}$ and isochrone no. 3 for higher masses, and assuming an age of ~ 1 Gyr. We checked that the change in results is negligible if the age is varied by a factor of up to 10. The typical uncertainty associated with the KIC T_{eff} values is 200 K, which translates to a uncertainty in mass of somewhat less than $0.1 M_{\odot}$. Vertical features in Figure 1, such as the gaps at $\sim 0.55 M_{\odot}$ and $\sim 0.7 M_{\odot}$, are artifacts introduced by the KIC temperature information and are not real. Conversion between $B-V$ and T_{eff} where required in this work was performed using the equations of Sekiguchi & Fukugita (2000).

The period measurements presented in this work are consistent with the existing ground-based photometric rotation period data, showing a trend of typically increasing rotation period with decreasing mass. The Sun, marked in Figure 1 as a red star, sits on the upper envelope period distribution.

The bimodality in period distribution, first reported by McQuillan et al. (2013a) for the M-dwarf sample, is clearly visible in the low-mass half of Figure 1. To explore the bimodality further, we plotted the data as a set of histograms, which are shown in Figure 2. To eliminate conversion uncertainties between T_{eff} and mass, we plotted the T_{eff} -period distribution in this figure. Each histogram is normalized, such that only frequencies on the period scale can be directly compared,

and not on the T_{eff} scale.

This histogram representation increases the clarity of the period bimodality in the low- T_{eff} region, and the width of the gap between the two sequences can be seen to increase towards cooler temperatures. At ~ 3500 K the two peaks are at ~ 20 days and ~ 40 days, with a Hartigan’s dip test (Hartigan & Hartigan 1985) p-value for unimodality of 0.01. At ~ 4000 K the two peaks are at ~ 14 days and ~ 30 days, with a Hartigan’s dip test p-value for unimodality of 0.15. This bimodal sequence is not visible above ~ 4500 K.

Figure 2 also shows that the upper envelope of periods increases steadily with decreasing temperature, from the hottest stars down to ~ 4500 K, at which point the long-period envelope decrease slightly before rising again below ~ 4000 K.

4.2. Amplitude of Periodic Variability

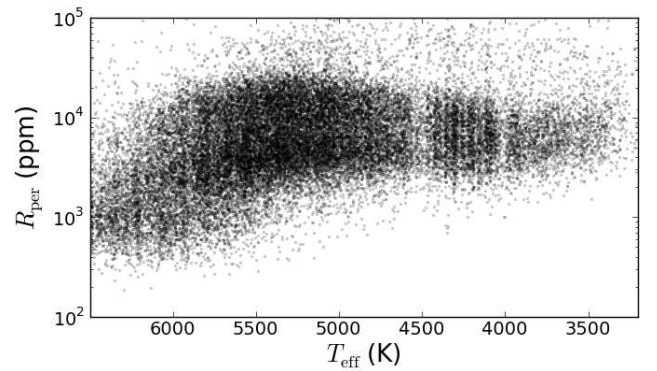


FIG. 3.— The distribution of periodic photometric variability amplitude, R_{per} , over the temperature range for all stars with derived rotation periods.

For each LC, we derived the amplitude of the photometric variability. We define this as a measure of the range between the 5th and 95th percentile of normalized flux, as described in Basri et al. (2011) and McQuillan et al. (2013a). To ensure we measure the amplitude of the periodic variability and not long-term trends in the data, we divide each LC into sections of period length and measure the 5th to 95th percentile of

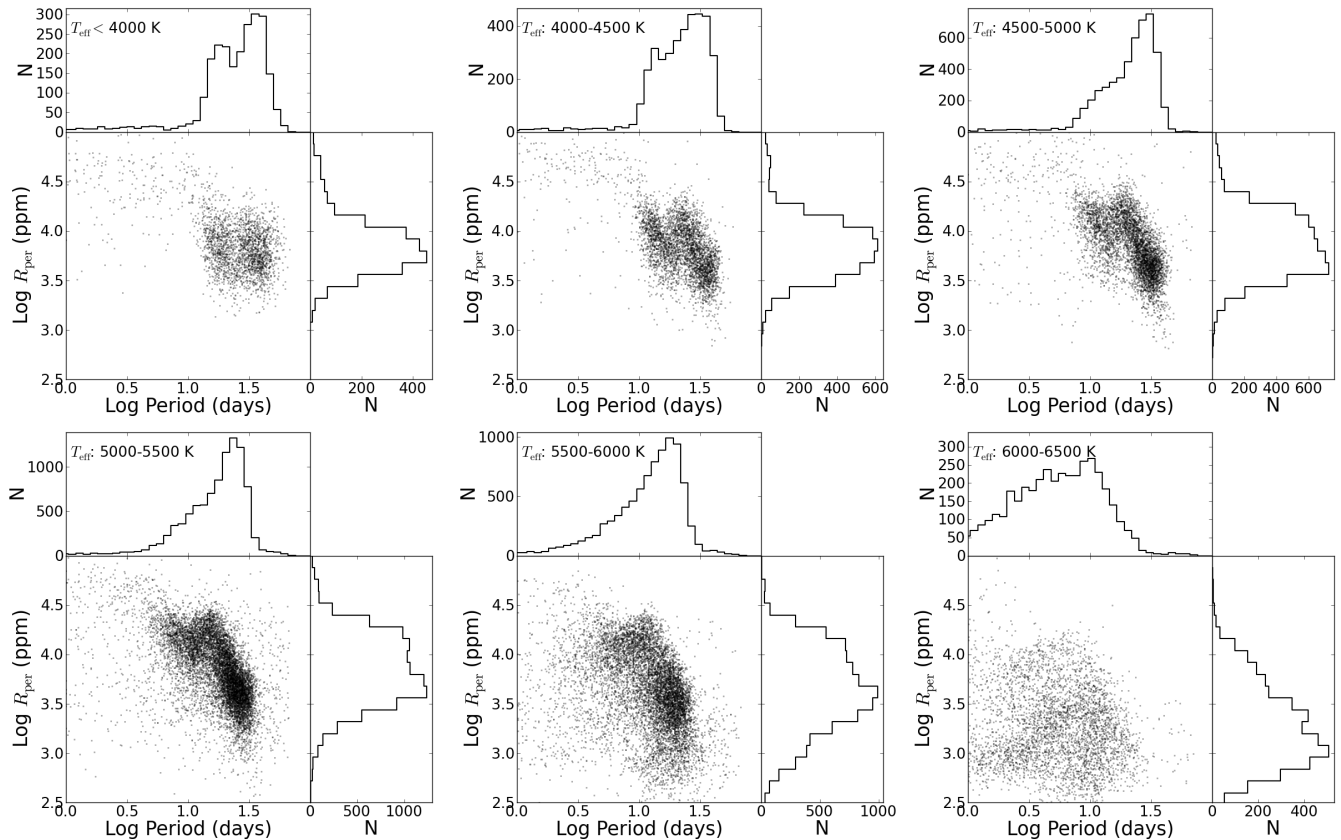


FIG. 4.— Log amplitude of periodic variability, R_{per} , against log period (for periods > 1 day), for each 500 K T_{eff} bin, as described in the top section of each panel. The top section of each panel shows the histogram of log period and the side section of each panel shows the histogram of log R_{per} .

normalized flux in each section. We then take the median of these values as the amplitude of periodic variability, R_{per} , for the LC.

Figure 3 shows R_{per} versus T_{eff} for the stars with detected rotation periods, and demonstrates a dearth of low-amplitude modulation at low temperatures. The range of amplitudes increases towards higher temperatures, with both an increase in the upper envelope of amplitudes and an excess of low-amplitude stars below ~ 5400 K. The typical range of amplitudes is between ~ 950 ppm (5th percentile) and $\sim 22,700$ ppm (95th percentile), with a median of $\sim 5,600$ ppm. We find typically higher amplitudes for shorter periods and lower effective temperatures.

We checked to see whether photon noise was artificially increasing the amplitude for fainter stars by examining the same distribution for only the brighter targets (Kepler magnitude < 14), and saw the same trend. The dearth of low-amplitude stars at low temperatures has been previously observed (Basri et al. 2010; McQuillan et al. 2012), but this is the first study to isolate variability resulting from rotational modulation. Since this figure shows only the stars for which rotation periods were detected, it suffers from the same biases as the period detection, namely that the long-period, low-amplitude stars are hardest to detect.

We also examined the distribution of photometric variability as a function of period, which is shown across the range of temperatures in Figure 4. Previous studies have suggested a typical increase in variability with

rotation rate (e.g., Pizzolato et al. 2003; Hartman et al. 2011), although this was not seen by Basri et al. (2011). In each T_{eff} bin a negative correlation between rotation period and R_{per} is evident, although the picture becomes somewhat more complex for the hottest stars. The cooler stars show a sequence running from high-amplitude, short-period stars, down to low-amplitude, long-period stars. This effect is unlikely to arise as a result of detection bias, since period detection is hardest for low amplitudes and long periods, and thus would not exclude the scarcely populated regions on either side of the sequence (low-amplitude and short-period, and high-amplitude and long-period).

The two period groups, shown clearly by the bimodality in the coolest T_{eff} bin, are visible up to 5500 K. The period-amplitude gradient prevents the bimodality from appearing in the period histograms, except for in the coolest set where it is most prevalent. It is interesting to note that the amplitudes of the short- and long-period groups in the coolest stars have similar amplitudes, whereas slower rotators extend to lower amplitudes for the early K and G stars.

5. GYROCHRONOLOGY

In Figure 5 we compare the T_{eff} -period sequence to the commonly used empirical gyrochronology models of Barnes (2007), Mamajek & Hillenbrand (2008) and Meibom et al. (2009). We find that the upper envelope of the period distribution is broadly consistent with a gyrochronological age of 4.5 Gyrs. The models all appear

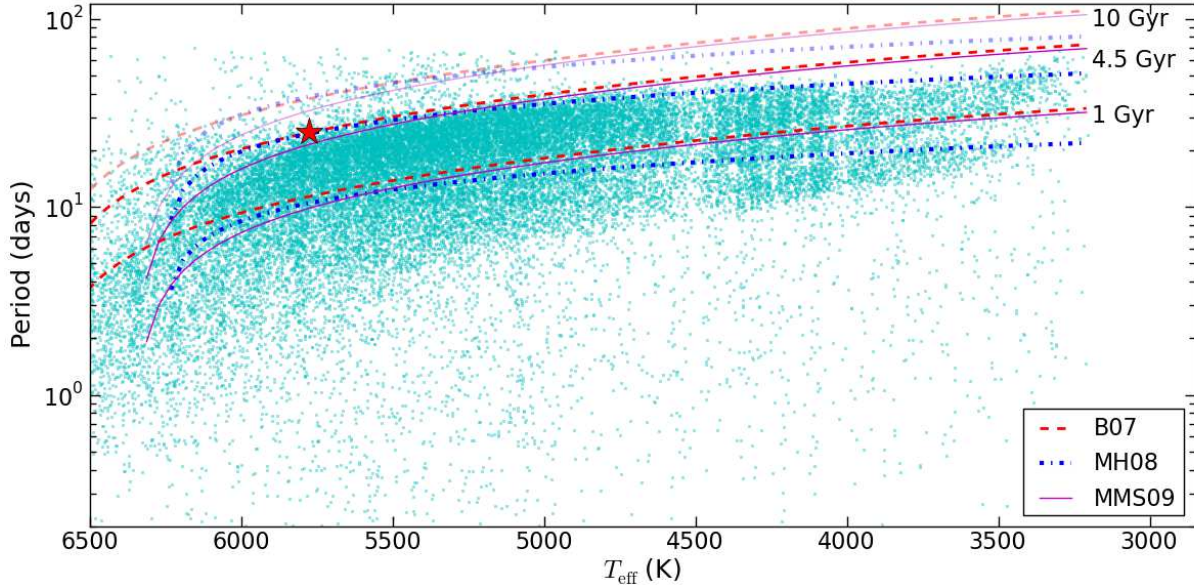


FIG. 5.— The T_{eff} -period distribution showing the location of selected rotational isochrones. The red dashed line marks the 1 Gyr (lower), 4.5 Gyr (middle) and 10 Gyr (upper, faint) empirical gyrochronology relation of Barnes (2007), denoted B07. The blue dot-dashed line marks the equivalent age isochrones from the empirical gyrochronology relation of Mamajek & Hillenbrand (2008), denoted MH08, and the solid magenta line marks that of Meibom et al. (2009), denoted MMS09.

to under-predict stellar ages, since they imply a significant proportion of the stars are < 1 Gyr and very few are > 4.5 Gyr, which would be unlikely for a field star population such as that observed by *Kepler*. Although the spin-down rate reduces with age, stars older than 4.5 Gyr should still appear distinctly above the upper envelope of points, as shown by the 10 Gyr isochrone. However, the observational biases are not well understood and the oldest stars are the hardest to derive photometric rotation periods for since they have long rotation periods and low-amplitude variability.

None of the models capture the slight dip in the upper envelope of rotation periods around 3600–3900 K, or the more pronounced increase in rotation periods for the M-dwarfs. None of the models are able to accurately reproduce the shape of the lower envelope of the distribution either, which should be less affected by observational or measurement biases than the upper envelope.

There is also a subset of very rapid rotators across all temperatures with periods below ~ 15 days, implying very young ages based on the gyrochronology models. This sample of very short-period rotators is considerably larger than would be expected for the *Kepler* sample, which are predominantly field stars and should not have a large number of stars younger than 1 Gyr. It is therefore unlikely that these objects are fast rotators due to youth. Possible alternative explanations include the hypothesis that these are non-eclipsing binaries, where spin-orbit interaction has lead to the rapid rotation speed, as suggested by McQuillan et al. (2013a), contaminating pulsation measurements, or that the more massive of these stars may represent the sub-giant contamination of the *Kepler* sample (van Saders & Pinsonneault 2013).

6. COMPARISON WITH PREVIOUS KEPLER ROTATION STUDIES

During the course of this work, two rotation period samples from the *Kepler* data were published by

Reinhold et al. (2013) and Nielsen et al. (2013). In this section we summarize their methods and compare their results to that of AutoACF.

6.1. Method of Reinhold et al. (2013)

Reinhold et al. (2013) used Q3 PDC-MAP data and selected stars with $\log g > 3.5$ and variability amplitude $R_{\text{varB}} \geq 0.003$, where R_{varB} (Basri et al. 2010, 2011) is the 5th and 95th percentile of normalized flux, with 4 hr boxcar smoothing of the LC. They applied the Generalized Lomb-Scargle periodogram (Zechmeister & Kürster 2009) to LCs binned to 2 hr sampling, in order to reduce computation time, and accepted periods in the range 12 hrs–45 days. Periods close to the orbital periods of known binaries were also excluded.

To minimize the aliasing caused by active regions on opposite hemispheres of the star (discussed in more detail in McQuillan et al. (2013a)), they compared the first two periods detected by the global sine fit, and if the difference between the longer period and twice the shorter period was less than 5%, then the longer period was selected.

An automatic filter was applied based on the number of zero crossings per period in the smoothed LC, under the assumption that real stellar rotation signals should have a low number of zero crossings, whereas pulsations and irregular variations will have a higher number. In total, 24,124 rotation periods were accepted by these filters.

6.2. Method of Nielsen et al. (2013)

Nielsen et al. (2013) used Q2–Q9 of the PDC-MAP and applied a Lomb-Scargle periodogram (Frandsen et al. 1995). Known EBs and planet candidates were excluded from their sample, in addition to stars with $\log g < 3.5$. They considered a periodogram peak to be valid if it is greater than 4 times the white noise estimated from the root mean square (RMS) of the time series, and accepted periods in the range 1–30 days.

TABLE 4

A NUMERICAL COMPARISON OF AUTOACF, REINHOLD ET AL. (2013) AND NIELSEN ET AL. (2013). THE TOP ROW CONTAINS THE TOTAL NUMBER OF REPORTED ROTATION PERIOD MEASUREMENTS FROM EACH WORK. THE NEXT THREE ROWS SHOW THE NUMBER IN BOTH THE COLUMN AND ROW LABELLED WORK.

	AutoACF	Reinhold et al. (2013)	Nielsen et al. (2013)
Total Number	34,030	24,124	12,515
AND AutoACF	–	20,009	10,381
AND Reinhold et al. (2013)	20,009	–	9,292
AND Nielsen et al. (2013)	10,381	9,292	–

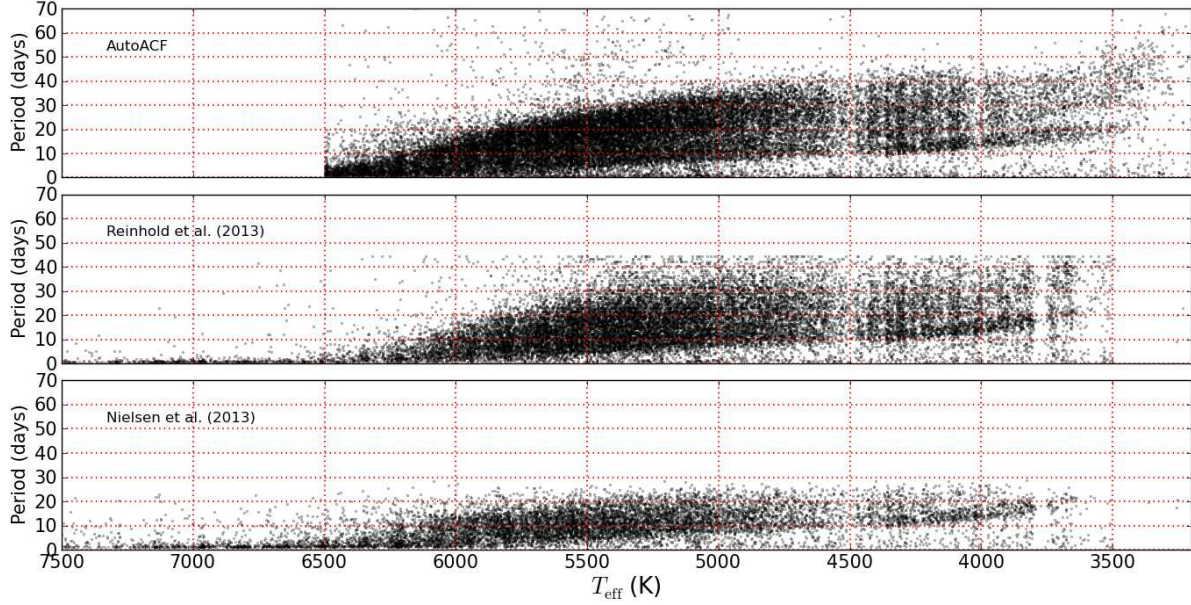


FIG. 6.— The T_{eff} -period distributions of AutoACF (top), Reinhold et al. (2013) (middle) and Nielsen et al. (2013) (bottom). A discussion of the comparisons between the different methods can be found in Section 6.3.

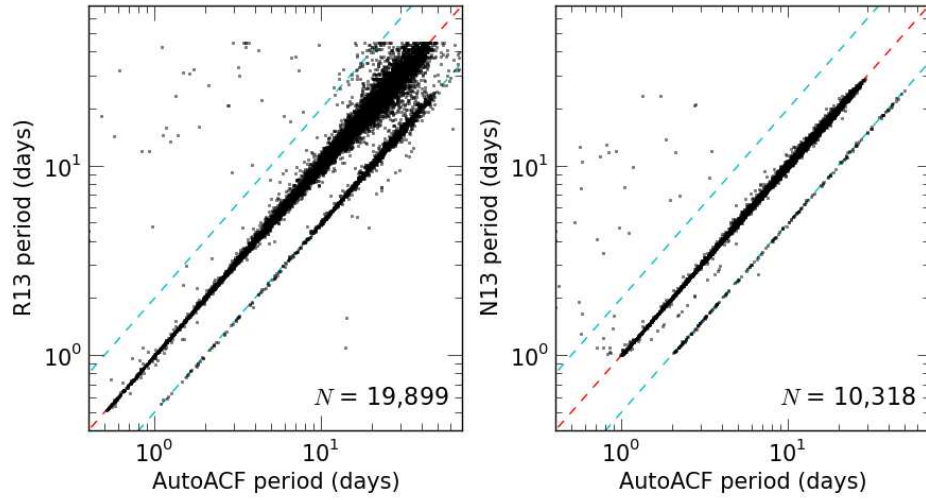


FIG. 7.— Direct comparisons between the periods measured in this work using AutoACF, and those detected by other groups. The panels show AutoACF compared to Reinhold et al. (2013) (R13, left) and Nielsen et al. (2013) (N13, right), with the number of stars in each group denoted as N . The red dashed line marks 1:1 ratio and the blue dashed lines mark 1:2 and 2:1.

Their selection routine required periods to be stable across multiple quarters, and the rotation period was then taken to be the median of the stable periods. They derived rotation periods for a total of 12,151 stars.

6.3. Comparison with AutoACF

A comparison between the number of rotation period detections by each method is displayed in Table 4, showing that AutoACF detects significantly more periods than the alternative methods. In some cases, AutoACF did not detect a period and the other methods did. This occurred in 4,115 cases when compared to Reinhold et al. (2013) and 1,770 compared to Nielsen et al. (2013). However, they had different initial sample selection criteria and when we compare only the stars to which AutoACF was applied, these numbers fall to 1,503 and 259 respectively. The cases missed by AutoACF with period detections from one of the other methods are predominantly those which failed to show coherent periods in multiple segments of the LC, and the remainder have low LPHs and fall just below our w_{thres} .

To compare our results to those of Reinhold et al. (2013) and Nielsen et al. (2013) we plotted period versus T_{eff} for each, which is shown in Figure 6. A similar structure is displayed in all three cases, with AutoACF extending to longer periods. All methods find a sequence of increasing period with decreasing temperature, with a small fraction of fast rotators at all temperatures.

The bimodality is clear in both the AutoACF and Reinhold et al. (2013) plots, and would likely have been found also by Nielsen et al. (2013) if longer periods were studied. Slight differences in the distribution below ~ 4000 K, most notably the vertical artifacts and lower T_{eff} range, occur as a result of our use of the Dressing & Charbonneau (2013) parameters in place of the KIC values, where available.

To directly compare our periods with those detected by Reinhold et al. (2013) and Nielsen et al. (2013), for the stars where both methods report periods, we plotted their periods versus ours in Figure 7. We find that in 13,035 and 9,757 cases respectively, the AutoACF period is within $0.05 P$ of that found by the alternative method, where P is the AutoACF period. However, it is immediately clear that in some cases, both alternative methods detect half the period reported by AutoACF. The periods of Reinhold et al. (2013) are within $0.05 (P/2)$ of half the AutoACF period in 1,304 cases (6.5% of their periods), and the same is true for the periods of Nielsen et al. (2013) in 202 cases (1.9% of their periods). We manually checked that in these cases, the AutoACF reflects the visually confirmed period and the alternative methods have detected incorrect periods. We note that for periods below 20 days, there is no evidence for AutoACF detecting half the period found by the other methods, however, for a small set of longer-period stars, Reinhold et al. (2013) find twice the AutoACF period.

For the alternative methods there are ~ 85 cases where AutoACF detects a short period and Reinhold et al. (2013) or Nielsen et al. (2013) detect a considerably longer period. These cases are predominantly caused by multiple periods in the LC, where AutoACF has selected the short period and the alternative methods have selected the longer period. The rest appear to be false detections by the alternative methods, most often arising

due to the short baseline used by Reinhold et al. (2013).

7. DISCUSSION

The present catalog, of more than 30,000 rotation periods and photometric amplitudes, was derived using a systematic automated search from a well defined sample of more than 133,000 main-sequence stars observed by *Kepler*, all of which have stellar temperature estimates. This sample can help us study the mass-age-period relation, and better understand its implication on the amplitude of photometric variability.

One clear feature of the sample is the fraction of rotation period detections as a function of temperature, which changes dramatically, from $\sim 80\%$ for the cool stars to $\sim 20\%$ for the hot stars of the sample. It seems that the photometric amplitude of the stellar rotation signature for the majority of the hot stars in the sample is smaller than our detection limit. This is supported by Figure 3, which shows that the detected amplitudes for the hot stars have a range that goes down to ~ 300 ppm, below which we do not detect periodicity. For the cool stars, on the other hand, the lower limit of the amplitudes is ~ 2000 ppm, well above our detection threshold. We might suggest that the incompleteness of our detections for the cool stars is therefore related to the orientation angle of the stellar axis of rotation relative to our line of sight, which prevents us from detecting the rotation of all cool stars in the sample.

Obviously, the photometric amplitude also carries information on the stellar surface inhomogeneity, presumably associated with the stellar magnetic field. Small amplitudes of variability can either result from smaller star spots or from a more even coverage of the stellar surface by spots. Therefore, our sample suggests that the typical inhomogeneity decreases as a function of stellar surface temperature. Further, Figure 4 shows that even for a given stellar temperature, the photometric amplitude is a decreasing function of the rotation period. This implies that the stellar inhomogeneity, probably together with the magnetic field, is decreasing as the star is aging, if we adopt the stellar rotation period as an age indicator.

The catalog shows that the typical rotation period increases with decreasing stellar mass and temperature. A comparison to empirical gyrochronology models (Barnes 2007; Mamajek & Hillenbrand 2008; Meibom et al. 2009) shows the upper envelope of points is broadly consistent with an age of 4.5 Gyrs, but that the models appear to under-predict the stellar ages, with a significant proportion falling below the 1 Gyr rotational isochrone. The position of the Sun close to the upper envelope is interesting, since this implies a lack of stars older than the Sun. More slowly rotating stars should still appear on the main sequence and would not have been removed by our exclusion of evolved stars.

The ‘dip’ feature in the upper envelope around $\sim 0.5 M_{\odot}$ (~ 4000 K), mentioned in McQuillan et al. (2013a) is still visible in the full sample. This mass corresponds to the transition point in young clusters, where lower mass stars spin up considerably more than stars above this mass (Henderson & Stassun 2012). This feature is not accounted for by any current gyrochronology models.

Thorough testing of AutoACF with synthetic LCs containing a range of periods and amplitudes is still required

to understand the effect of observational biases on the upper envelope, since low-amplitude, long-period signals are most strongly affected by systematics and quarter joining.

A series of period histograms across varying T_{eff} bins, shown in Figure 2, displays some of the finer details of the period distribution, including a significant bimodality in the low-mass stars, which becomes less apparent towards higher masses. This bimodality, forming two sequences in the mass- or T_{eff} -period distribution, cannot be explained by any observational or measurement biases, and thus requires a physical explanation. One hypothesis presented in McQuillan et al. (2013a) is that this is a possible hint of two epochs of star formation in the solar neighborhood. This effect may become less apparent towards higher masses, since these brighter stars are seen to greater distances, hence probing a larger region of space.

An alternative explanation, also presented in McQuillan et al. (2013a), is that the dearth of stars between the two sequences represents a mass-dependent, rapid transition phase in the stellar spin-down law, where few stars are observed. In this scenario the star formation rate is unimodal but the spin-down occurs as two slow stages, with a rapid stage in between, producing a bimodal distribution of periods. Both

these hypotheses require further follow-up observations and detailed modeling (e.g., Gallet & Bouvier 2013; Epstein & Pinsonneault 2014), which will be aided by the larger sample size and wider T_{eff} coverage of the new data.

Acknowledgments: The authors wish to thank Jerome Bouvier, Ariele Königl, Jennifer Van Saders and Jason Steffen for insightful discussions on this work. The research leading to these results has received funding from the European Research Council under the EU's Seventh Framework Programme (FP7/(2007-2013)/ERC Grant Agreement No. 291352). T.M. also acknowledges support from the Israel Science Foundation (grant No. 1423/11) and the Israeli Centers of Research Excellence (I-CORE, grant No. 1829/12). S.A. acknowledges support from STFC Consolidated grant ref. ST/K00106X/1 and Leverhulme Research Project grant RPG-2012-661. All of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts.

REFERENCES

- Affer, L., Micela, G., Favata, F., & Flaccomio, E. 2012, *MNRAS*, 424, 11
- Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, *PASP*, 125, 989
- Baliunas, S., Sokoloff, D., & Soon, W. 1996, *ApJL*, 457, L99
- Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, *A&A*, 337, 403
- Barnes, S. A. 2003, *ApJ*, 586, 464
- . 2007, *ApJ*, 669, 1167
- . 2010, *ApJ*, 722, 222
- Basri, G., Walkowicz, L. M., Batalha, N., et al. 2010, *ApJL*, 713, L155
- . 2011, *AJ*, 141, 20
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977
- Bouvier, J. 2013, in *EAS Publications Series*, Vol. 62, *EAS Publications Series*, 143–168
- Bouvier, J., Forestini, M., & Allain, S. 1997, *A&A*, 326, 1023
- Ciardi, D. R., von Braun, K., Bryden, G., et al. 2011, *AJ*, 141, 108
- Collier Cameron, A., Davidson, V. A., Hebb, L., et al. 2009, *MNRAS*, 400, 451
- Dressing, C. D., & Charbonneau, D. 2013, *ApJ*, 767, 95
- Epstein, C. R., & Pinsonneault, M. H. 2014, *ApJ*, 780, 159
- Frandsen, S., Jones, A., Kjeldsen, H., et al. 1995, *A&A*, 301, 123
- Gallet, F., & Bouvier, J. 2013, *A&A*, 556, A36
- Gizon, L. 2002, *Astronomische Nachrichten*, 323, 251
- Gizon, L., & Solanki, S. K. 2004, *Sol. Phys.*, 220, 169
- Goulding, N. T., Barnes, J. R., Pinfield, D. J., et al. 2012, *MNRAS*, 427, 3358
- Harrison, T. E., Coughlin, J. L., Ule, N. M., & López-Morales, M. 2012, *AJ*, 143, 4
- Hartigan, J. A., & Hartigan, P. M. 1985, *The Annals of Statistics*, 13, 70
- Hartman, J. D., Bakos, G. Á., Noyes, R. W., et al. 2011, *AJ*, 141, 166
- Henderson, C. B., & Stassun, K. G. 2012, *ApJ*, 747, 51
- Huber, D., Silva Aguirre, V., Matthews, J. M., et al. 2014, *ApJS*, 211, 2
- Irwin, J., Aigrain, S., Bouvier, J., et al. 2009, *MNRAS*, 392, 1456
- Irwin, J., Berta, Z. K., Burke, C. J., et al. 2011, *ApJ*, 727, 56
- Irwin, J., & Bouvier, J. 2009, in *IAU Symp.*, Vol. 258, *The Ages of Stars*, 363–374
- Kaler, J. B. 1989, *Stars and their spectra. an introduction to spectral sequence* (Cambridge: Cambridge University Press)
- Kawaler, S. D. 1988, *ApJ*, 333, 236
- . 1989, *ApJ*, 343, L65
- Kenyon, S. J., & Hartmann, L. 1995, *ApJS*, 101, 117
- Kiraga, M., & Stepien, K. 2007, *ACTAA*, 57, 149
- Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, *ApJ*, 713, L79
- Mamajek, E. 2011, *A Modern Mean Stellar Color and Effective Temperatures (Teff) Sequence for O9V-Y0V Dwarf Stars*, Online: accessed 24-Feb-2013
- Mamajek, E. E., & Hillenbrand, L. A. 2008, *ApJ*, 687, 1264
- Matt, S. P., MacGregor, K. B., Pinsonneault, M. H., & Greene, T. P. 2012, *ApJ*, 754, L26
- McQuillan, A., Aigrain, S., & Roberts, S. 2012, *A&A*, 539, A137
- McQuillan, A., Aigrain, S., & Mazeh, T. 2013a, *MNRAS*, 432, 1203
- McQuillan, A., Mazeh, T., & Aigrain, S. 2013b, *ApJ*, 775, L11
- Meibom, S., Mathieu, R. D., & Stassun, K. G. 2009, *ApJ*, 695, 679
- Meibom, S., Mathieu, R. D., Stassun, K. G., Liebesny, P., & Saar, S. H. 2011, *ApJ*, 733, 115
- Moraux, E., Artemenko, S., Bouvier, J., et al. 2013, *A&A*, 560, A13
- Nielsen, M. B., Gizon, L., Schunker, H., & Karoff, C. 2013, *A&A*, 557, L10
- Pizzolato, N., Maggio, A., Micela, G., Sciortino, S., & Ventura, P. 2003, *A&A*, 397, 147
- Prša, A., Batalha, N., Slawson, R. W., et al. 2011, *AJ*, 141, 83
- Reiners, A., & Mohanty, S. 2012, *ApJ*, 746, 43
- Reinhold, T., Reiners, A., & Basri, G. 2013, *A&A*, 560, A4
- Sekiguchi, M., & Fukugita, M. 2000, *AJ*, 120, 1072
- Slawson, R. W., Prša, A., Welsh, W. F., et al. 2011, *AJ*, 142, 160
- Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, *PASP*, 124, 1000
- Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. 2012, *PASP*, 124, 985
- van Saders, J. L., & Pinsonneault, M. H. 2013, *ApJ*, 776, 67
- Walkowicz, L. M., & Basri, G. S. 2013, *MNRAS*, 436, 1883
- Zechmeister, M., & Kürster, M. 2009, *A&A*, 496, 577

APPENDIX

SECTION A: AUTOMATION OF THE ACF

Our previous work using the ACF method has relied upon visual verification of the period detection of each LC (McQuillan et al. 2013a,b). The large number of stars in the full *Kepler* sample makes this approach unfeasible, and therefore we opted to design an automated period verification procedure. In this appendix we describe the design and testing of the automated ACF (AutoACF) method.

Visually Examined Training Set

A large subset of *Kepler* targets were visually examined to form a training set for the automation procedure. During this process, each visually examined target was assigned a classification:

1. Periodic: Clear periodicity in LC, with the selected ACF peak matching the visually identified period.
2. Wrong Peak: Clear periodicity in LC but incorrect peak selected in ACF (usually half or twice the rotation period identified by eye).
3. Probable Periodic: Some indication in the LC and ACF that the star is periodic, but not clear enough to include in the ‘Periodic’ training set.
4. Likely Pulsation: The LC and ACF indicate that the periodicity detected is more likely to arise from effects other than rotation. See Section 3.5 of McQuillan et al. (2013a) for a more detailed discussion.
5. Non Periodic: The LC and ACF show no signs of periodicity, or the ACF peak clearly results from instrumental systematics.

Figure 8 shows examples of these 5 classes: The first column shows a clear case of periodicity; the second column shows an example where the rotation period can be identified by eye as ~ 30 days, but due to residual systematics in the LC, the second ACF peak is higher and therefore erroneously selected to represent the period; the third column shows a case which is likely to be periodic based on the repeated features in the LC and the peaks in the ACF, but is not clear enough to make it into the ‘Period’ set; the fourth column shows rapid and irregular variability which gives rise to irregular peaks in the ACF, which we consider unlikely to arise from rotational modulation; and the final column shows a case with a noise-dominated LC and no obvious peaks in the ACF.

We divided the sample into 500 K T_{eff} bins, and visually examined the LCs and ACFs in a random order until 1000 verified periods (categories 1 and 2) were classified in each bin. The results of the above identification process are shown in Table 5.

At this stage we tested a variety of statistics to distinguish automatically between the visually confirmed detection and the false positives. This included using the height, width and number of the detected ACF peaks, the uncertainty on the period measurements, and the stability of the period detection over different quarters. Ultimately, we converged on the two-stage selection method described below.

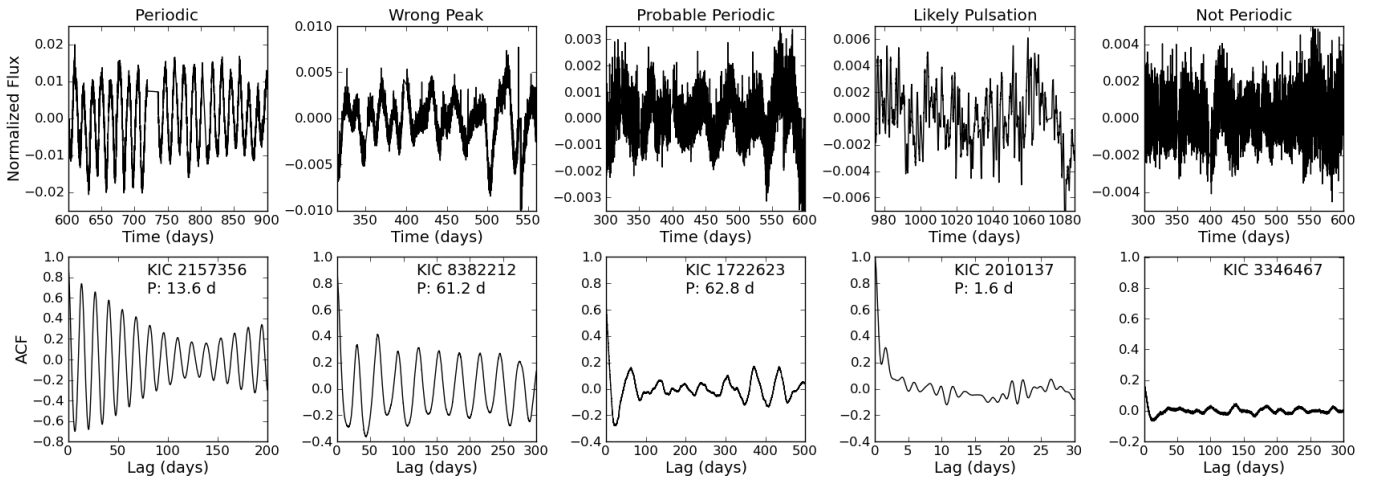


FIG. 8.— Example LCs (top) and ACFs (bottom) of the different classifications listed in Section A.1, with the classification type noted above each column.

TABLE 5

TRAINING SET SELECTION: TEMPERATURE BIN, TOTAL NUMBER OF TARGETS, NUMBER VISUALLY EXAMINED TARGETS, THE VISUAL CLASSIFICATIONS DESCRIBED IN THE TEXT, AND THE PERIODIC FRACTION OF THE VISUALLY EXAMINED SET.

T_{eff}	Total	Visually Examined	Periodic	Wrong Peak	Probable Periodic	Pulsation	Non Periodic	Periodic %
3500–4000	1657	1657	1342	26	181	47	61	81.0
4000–4500	3875	1466	994	6	293	31	142	68.2
4500–5000	7761	2105	987	13	417	75	613	47.5
5000–5500	23767	3337	988	12	737	64	1536	30.0
5500–6000	37839	4580	985	15	797	74	2709	21.8
6000–6500	14565	3644	978	22	589	82	1973	27.4
Total	89464	16789	6274	94	3014	373	7034	37.4

Selection Process 1 - Coherent Periods in Multiple LC Segments

The separation of periodic and non-periodic ACF results was performed in two steps. A real astrophysical periodicity is more likely to be detected in multiple regions of the LC than a systematic effect or artifact would. Our first selection procedure was therefore to accept only ACF detections that are consistent in different parts of the LC. We adopted the criterion that the period detected in the full Q3–Q14 LC must be coherent with that found in 2 out of the 4 segments of the data, where a segment is defined as 3 consecutive *Kepler* quarters (Q3–Q5, Q6–Q8, Q9–Q11, Q12–Q14). The period detection in a segment is defined as coherent if it is within 10% of the period (or half / twice the period) detected in the full Q3–Q14 LC.

This method selected 63.8% of the training sample, of which 63.9% were classed as ‘Periodic’. This selection retained 99.3% of the ‘Periodic’ classifications from the training set. The multiple segment criterion is not sufficient alone to distinguish between true and false period detections. Due to the rotation of the *Kepler* satellite through 90° between quarters, the systematics can repeat every 4 quarters when targets appear in the same place on the CCD. This can lead to matching, but false period detections in multiple segments. It is also possible to have low-amplitude ACF peaks arising from noise, which are consistent by coincidence over multiple segments. We therefore required a second selection stage.

Selection Process 2 - ACF Local Peak Height Weighting

The local peak height (LPH) of the ACF is defined as the height of the selected peak with respect to the mean of the troughs on either side. Visually confirmed period detections typically have higher LPHs than those suspected to originate from noise or systematic effects. The LPH is also correlated with period, where shorter periods have higher LPHs. This is a result of the increased effect of instrumental systematics and LC processing on longer periods.

To exploit the properties of LPH for distinguishing between periodic and non-periodic ACFs we assign a weight, w , between 0–1, based on LPH, period and T_{eff} , where a higher weight indicates a greater probability of a true period detection. This weight comprises two parts, which are described below and shown in Figure 9. A threshold, w_{thres} can then be applied to distinguish between periodic and false detections. T_{eff} is incorporated into the selection criteria because the visually confirmed period distribution varies with temperature, and hence the probability of a detection being real at a certain period and LPH is different for targets of different T_{eff} .

Weight Part 1: Magnitude of LPH

The first part of the weight, w_1 , is dependent on the magnitude of LPH, adjusted by a positive gradient to account for the fact that at longer periods, higher LPHs are found for false detections than at short periods. If the LC has systematic effects in two or three quarters, these can easily produce long-period ACF peaks with high LPHs, whereas short period effects would require many more coherent repetitions to produce an ACF peak with the same LPH, and hence have a higher likelihood of being real detections. The baseline, marked as the dashed line in each panel of Figure 9, is described by the equation: $\text{LPH}_{\text{line}} = 0.0027P + 0.126$, where P is period. This line was found by fitting a straight line to the period of the ‘Not Periodic’ points as a function of LPH for the middle two temperature bins (4500–5500 K) and adding 0.05 to the y-offset to approximate the upper edge of points rather than the central fit:

The w_1 value for each LC is then calculated from

$$w_{1,i} = 1.0 - \exp[-(\text{LPH}_i - \text{LPH}_{\text{line}})], \quad (\text{A1})$$

where negative values of $(\text{LPH}_i - \text{LPH}_{\text{line}})$ are set to 0, and $i = 0, 1, 2, \dots, N$ where N is the number of LCs. This gives a sharp increase in w_1 at low LPHs and then a more gradual increase at high LPH, which is required to separate true detection from the continuum of false detections.

Weight Part 2: Position in T_{eff} -Period-LPH Space

The second part of the weighting, w_2 , is based on the location of the target in T_{eff} -period-LPH space. Since ‘Periodic’ detections typically lie along a sequence of decreasing LPH with increasing period (see Figure 9), detections further from this band are more likely to be false. This is especially important at low LPHs, where a ‘Periodic’ detection at the base of the periodic sequence can have the same LPH as a ‘Non Periodic’ detection further from the base of the sequence. We therefore identified the position and width of this sequence as a function of T_{eff} , and developed a criterion for selection based on distance from the expected location in T_{eff} -period-LPH space.

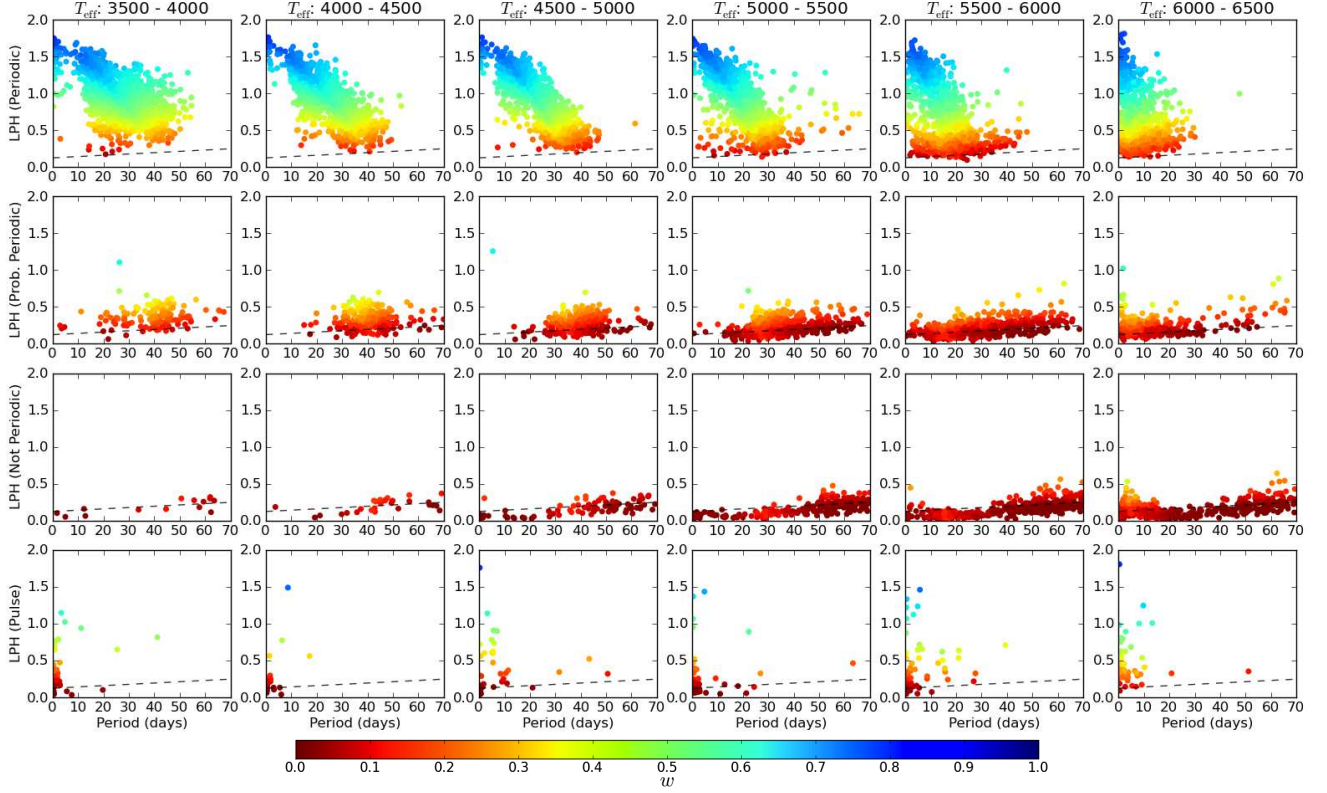


FIG. 9.— Each panel shows the period-LPH distribution, colored by total weight, w . The top row displays the ‘Periodic’ class in each T_{eff} bin, the second row shows the ‘Probable Periodic’ class, the third row is the ‘Non Periodic’ class and the bottom row is the ‘Pulsation’ class. The dashed line, described in the text, shows the baseline from which w_1 is measured.

The position and width, α , of the sequence were found using a polynomial fit to the median, 5th and 95th percentiles of the period at intervals in LPH for each temperature bin. We used linear interpolation between the temperature bins to obtain an expected period, P_{ex} , as a function of T_{eff} and LPH, and α as a function of T_{eff} .

The distance from this expected period, for a given LPH measurement is converted to w_2 by:

$$w_{2,i} = \frac{\exp(-(d_i/\alpha_i))}{\text{MAX}(\exp(-(2d_i/\alpha_i)))}, \quad (\text{A2})$$

where $d_i = |P_i - P_{i,\text{ex}}|$. By combining w_1 and w_2 , an overall weight w is obtained, which has a range between 0–1, with higher values at higher LPHs, with the exception of close to the expected period, where the weight is boosted even for relatively low LPHs. The weights combined as $w = (w_1 + Cw_2)/C$, where C is a parameter of the automation routine which is adjusted to give the maximum separation between detection and false positives from the training set. The resulting distribution of w is displayed in Figure 9, which shows the training set results. To avoid excess false positives from pulsators and unfeasibly long periods, we only consider periods in the range 0.2–70 days. This range was chosen by visual examination of the LCs in the training set.

The parameter C , which controls the ratio of the weight combination, and the threshold w_{thres} were selected to optimize the distinction between periodic and false positives. Our priority was primarily to minimize the number of false positives (‘Non Periodic’ class), while including the highest possible percentage of the ‘Periodic’ class and a minimal number of the ‘Probable Periodic’ class. By running AutoACF on the training set, using a range of possible values for C and w_{thres} , we found that $C = 0.15$ and $w_{\text{thres}} = 0.25$ produce the best match to our selection criteria. Tables 1 and 2 includes w so that users of this data may select their own threshold to be more or less conservative.

We considered the potential of the T_{eff} -dependent component of the weighting to bias the result towards periods that fit the global trend, but concluded that the effect will not preferentially exclude periods outside the expected trend region, provided their LPH is sufficient indicate a real detection. The second weighting is designed to select additional real periods which are found to fall below a simple LPH threshold, as indicated by visual inspection.

Training Set Results

Table 6 shows the outcome of AutoACF applied to the training set, and demonstrates that the method is able to select 86.2% of the visually identified ‘Periodic’ class with only a 0.1% false positive rate (‘Non Periodic’). The fractions of each class in the selected group represent the predicted fractions obtained when applying AutoACF to the full *Kepler* sample.

To explore the differences between the visually identified periodic set and the AutoACF selected members of the

TABLE 6

TRAINING SET RESULTS: SELECTED NUMBERS PER CLASSIFICATION, USING $w_{\text{crit}} = 0.25$. THE ‘SELECTED’ ROW SHOWS THE NUMBER OF EACH CLASS DETERMINED AS PERIODIC BY AUTOACF. THE PERCENTAGES IN BRACKETS ARE PERCENT OF THE SELECTED SETS CONSISTING OF EACH CLASS, I.E., 94.5% OF THE ‘SELECTED’ SET IS THE ‘PERIODIC’ CLASS.

	Periodic	Wrong Peak	Probable Periodic	Pulsation	Non Periodic	Total
Selected by AutoACF	5411 (94.5%)	32 (0.6%)	191 (3.3%)	82 (1.4%)	8 (0.1%)	5724

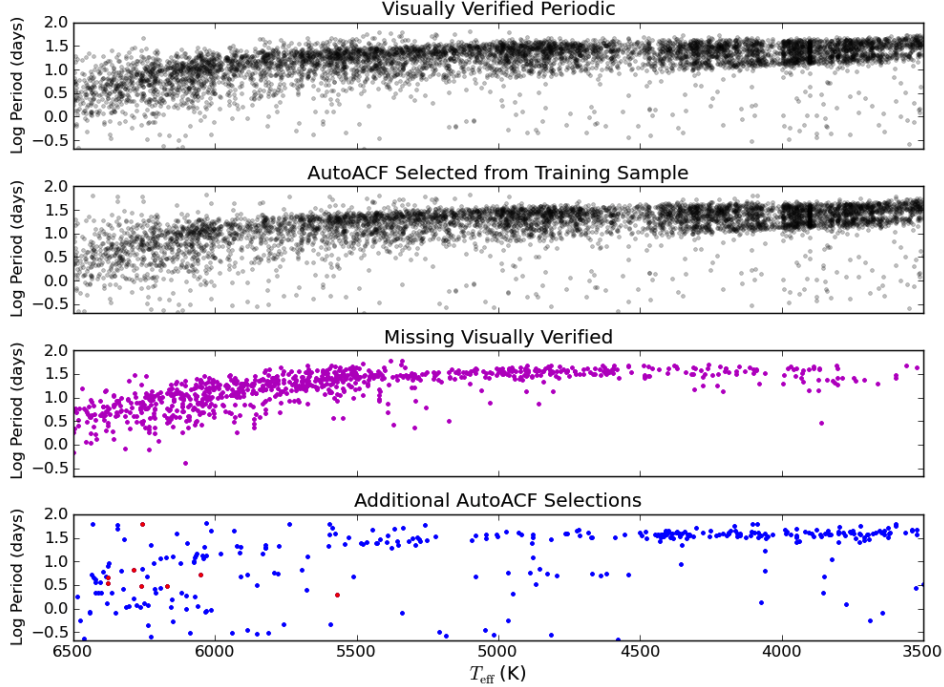


FIG. 10.— Training Set: Comparison between the visually examined ‘Periodic’ class and the AutoACF selection of the visually examined sample. Top panel: Visually verified ‘Periodic’ sample. Second panel: AutoACF selected sample. Third panel: The ‘Periodic’ class targets not in the AutoACF selection (magenta). Fourth panel: The ‘Probable Periodic’ (blue) and ‘Non Periodic’ (red) selected by AutoACF.

training set, we plotted comparisons in the T_{eff} -period distribution (Figure 10). In each panel the sequence populated by real rotation period detections is clearly shown. The first and second panels directly compare the visually confirmed ‘Periodic’ class and the AutoACF selected set, showing that the two differ very little from one another. The third and fourth panels display in color the detections that are different between the visual and AutoACF selection methods. The magenta points in the third panel are visually verified ‘Periodic’ detections that are not selected by AutoACF. The colored points in the fourth panel show AutoACF selected points that are not in the visually ‘Periodic’ class. These primarily consist of ‘Probable Periodic’ detections (cyan) with a small number of the ‘Non Periodic’ class (blue). The additional ‘Probable Periodic’ targets are typically long period detections, and their inclusion by AutoACF is expected, since longer period signals are more strongly affected by systematics and LC processing, which may cause the signal to be less clear to visual inspection. The differences between the visually inspected ‘Periodic’ sample and the AutoACF sample do not significantly alter the shape of the period- T_{eff} distribution.

SECTION B: ADDITIONAL DETECTIONS

The following LC and ACF corrections were applied to those targets not initially selected as periodic by AutoACF. Large faults in the LC, such as discontinuities and uncorrected quarters, can cause a steep underlying slope in the ACF, and while the ACF peaks are still visible by eye, the LPH is reduced to a level not selected by AutoACF. We therefore performed an automatic search to remove sections of data containing faults or uncorrected data. This was done by rejecting quarters where the range of normalized flux between the 5th and 95th percentile was more than 5 times the median range of the other quarters. Running AutoACF on the fault-corrected LCs resulted in an additional 56 period detections, marked with the flag ‘BQR’ in Table 1. An example LC and ACF before and after fault removal is shown in Figure 11.

In cases where more than one period are present in the LC, the ACF displays an underlying oscillation from the longest period, with small peaks superimposed on it for the shorter periods. These short period peaks in the ACF typically have a low LPHs due to the underlying large peaks from the longer periods, which lead to a rejection of the detection by AutoACF. To deal with cases of multiple periods in the LC, we smoothed the ACF to allow detection of the larger underlying peaks, instead of the short period oscillations superimposed on it. The smoothing was performed by convolving the ACF with a Gaussian kernel, of window size of either 60 data points and full width at half-maximum (FWHM) of 20 data points (soft smoothing) or a window size of 200 data points and FWHM of 100 data points (hard

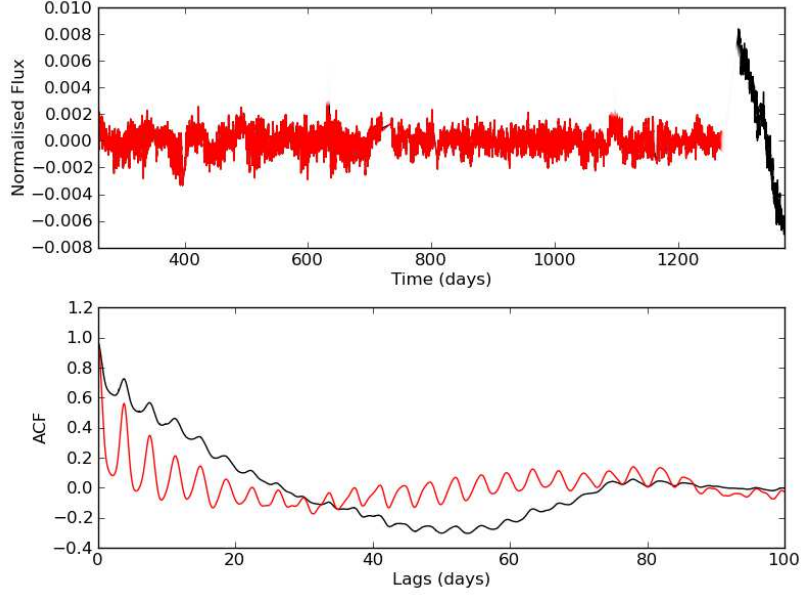


FIG. 11.— Example LC and ACF (KIC 7286309), where the PDC-MAP correction of a quarter has failed, resulting in a steep slope at short lags in the ACF. The original LC and ACF are shown in black and the corrected ones in red. After removal of the faulty data, the ACF displays clearer peaks with higher LPHs.

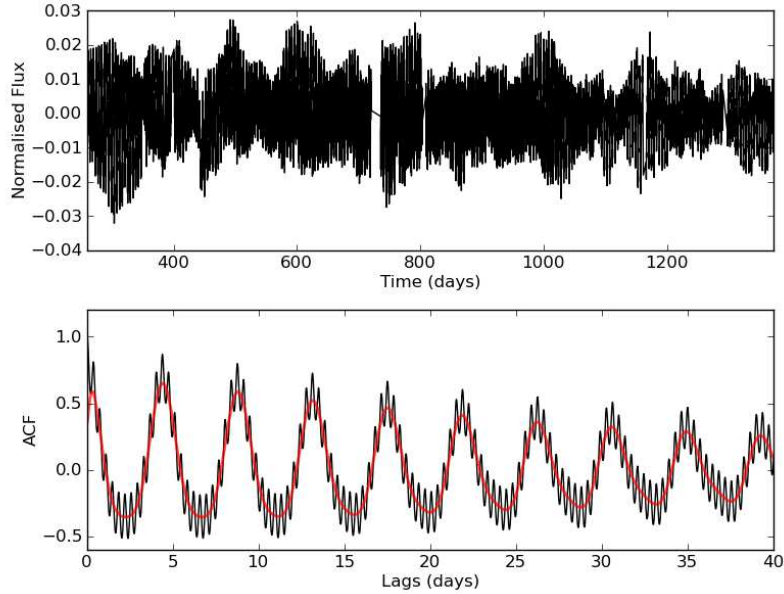


FIG. 12.— Example LC and ACF (KIC 6064171), which displays both high- and low-frequency periodicity. The smoothed ACF (shown in red) has clear, high LPH peaks, which can be detected by AutoACF.

smoothing).

We first applied the ‘soft smoothing’ to all targets for which no period was previously detected, and visually inspected the results of AutoACF. This provided an additional 253 period detections, marked with the flag ‘SM1’ in Table 1. We then ran the ‘hard smoothing’ on the remaining targets, and those identified by eye as requiring further smoothing from the ‘soft smoothing’ set. This process, followed by visual examination of the results, yielded an additional 125 rotation periods, marked with the flag ‘SM2’ in Table 1. An example LC and ACF for a smoothed case are shown in Figure 12. These additions bring the total number of rotation period detections to 34,030.